Journal of Marine Science and Technology

Volume 29 | Issue 3 Article 14

Investigating Model Solution Correctness for Parameter Uncertainty in Both Objective Function and Constraints

Shangyao Yan
Department of Civil Engineering, National Central University, Taoyuan, Taiwan, R.O.C., t320002@cc.ncu.edu.tw

Sin-Siang Wang
Department of Civil Engineering, National Central University, Taoyuan, Taiwan, R.O.C.

Chun-Yi Wang
Department of Civil Engineering, National Central University, Taoyuan, Taiwan, R.O.C.

Recommended Citation
Yan, Shangyao; Wang, Sin-Siang; and Wang, Chun-Yi (2021) "Investigating Model Solution Correctness for Parameter Uncertainty in Both Objective Function and Constraints," *Journal of Marine Science and Technology*: Vol. 29: Iss. 3, Article 14.
DOI: 10.51400/2709-6998.1468
Available at: https://jmstt.ntou.edu.tw/journal/vol29/iss3/14

RESEARCH ARTICLE

Investigating Model Solution Correctness for Parameter Uncertainty in Both Objective Function and Constraints

Shangyao Yan*, Sin-Siang Wang, Chun-Yi Wang

Department of Civil Engineering, National Central University, Taoyuan 32001, Taiwan, R.O.C

Abstract

Parameter uncertainty, which may arise due to changes in the environment or human error, may be incorporated into the objective function and the constraints in an optimization model. However, to simplify the modeling, the values of these parameters are usually set or projected as deterministic values. It is no wonder that the modelling results based on these inaccurate parameters are neither correct nor reliable. Thus, it is important to examine the correctness of the model results in relation to parameter uncertainty. This study aims to analyze solution correctness in relation to different degrees of parameter uncertainty for the parameters in the objective function and the constraints, specifically for a project scheduling model. To examine the relationship between the solution correctness, the parameter uncertainty and the solution tolerance error, we conduct a numerical experiment including a number of different scenarios, each associated with a degree of uncertainty for all parameters in both the objective function and the constraints. Finally, the regression technique is adopted to more efficiently analyze the relationship between model input error, solution tolerance error and model output error, by estimating equations representative of their relationship. The obtained results and findings could be useful for the planners to apply any optimization models, including maritime transport optimization models, and to design solution algorithms in practice.

Keywords: Solution correctness, Parameter uncertainty, Optimization model, Regression

1. Introduction

Optimization models are a good tool for decision makers to solve complicated optimization problems, such as maritime transport optimization problems, in practice. Based on the given parameters with the constraints and the objective for a problem, mathematical optimization applies an exact solution algorithm embedded in a complicated combinational analysis of variable values to choose the best solution from among numerous feasible combinations of variable values. Because all model inputs synthetically affect the model output, the effects of the parameters on the model output are extremely complicated.

To effectively apply an optimization model, related parameters are required. However, the parameters calculated or estimated in the real world may not be correct due to environmental stochasticity or human error. It is quite difficult to accurately calculate uncertain parameter values without errors. When uncertain parameters are applied in an optimization model, the solution may not be correct. The main reason for this is that the existence of model input errors will cause the occurrence of model output errors, which is the "blind spot" for the successful application of optimization models. Furthermore, the decision maker may not know that

Received 15 October 2020; revised 16 November 2020; accepted 18 November 2020.
Available online 25 June 2021.

* Corresponding author.
E-mail address: t320002@cc.ncu.edu.tw (S. Yan).

https://doi.org/10.514/2709-6998.1468

the "optimal" decision made when a model solution contains errors is inappropriate. Additionally, the gap between the solution obtained from an optimization model with uncertain parameters and the real optimal solution for the model is unknown. If the gap is significantly large, then the so-called optimal solution is not meaningful. If so, a solution obtained from an approximation solution algorithm with an acceptable solution tolerance error might be more useful than the solution gained from an optimization model with uncertain parameters. In the past, this type of situation has rarely been discussed, but further work is indeed needed to understand the practical applications of optimization models, and how to design better optimization models and approximation solution algorithms.

A number of methods have been used in the past to estimate suitable parameter values for optimization model applications. Methods adopted for evaluating uncertain parameters include grey prediction, data mining, robust optimization, artificial neural networks, machine learning, fuzzy sets, stochastic programming and so on. For instance, see Refs. [1–17]. However, the solutions obtained in past studies were mainly evaluated by comparing them with the best solutions found previously because the true optimal solutions cannot be obtained from models with uncertain parameters. In other words, the solution performance for such a model, even one obtained with any of the aforementioned approaches, cannot be objectively evaluated, because there are still errors in the effective solutions.

Variance-based sensitivity analysis has also been used for assessing the importance of model inputs when there are probability distributions associated with each input to quantify the relative contribution of uncertainty from different sources [18]. Moreover, the variance decomposition method, coupled with variance-based sensitivity analysis, has been proposed as a means to efficiently examine the effect of an individual model input on the model output with input independence [19]. However, sensitivity analysis has only been efficiently carried out for linear programming models with changes in one parameter, which are known as optimization models. Two types of parametric programming analysis for linear programming models have been developed to examine continuous decreasing/increasing changes in a set of parameters (e.g., see Hiller and Lieberman, [20]). No efficient approach has yet been developed for performing sensitivity analysis of linear programming models or integer programming models, with simultaneous changes in two or more parameters, such as Chen et al.'s [21] model.

There have been some studies that utilize approximation solution algorithms with a solution tolerance error to increase the solution efficiency of models that include uncertain parameters. For instance, see Refs. [22–30]. Although the solution efficiency obtained is good, the effect of various solution tolerance errors on model solutions with input errors has not been discussed. The complicated combinatorial analysis of variable values from an exact solution algorithm and a heuristic algorithm could result in a gap between the optimal solution and approximate solution. Furthermore, if the model parameters contain errors, then this gap will be distorted, and since the changes in both solutions are unknown, the superiority of the "optimal" solution over the "heuristic" solution may not hold. It is necessary to obtain a better understanding of the gap based on the given parameter uncertainty and solution tolerance error.

There are basically two types of factors causing uncertainty of parameters in a model: uncontrollable and controllable. Uncontrollable factors are the result of nonhuman behavior. For instance, discrepancies in information between the supply and demand sides are considered to be an uncontrollable factor. Such discrepancies will cause differences between the currently estimated uncertain parameter values and those estimated in the future, meaning that the estimated values will include errors. Random error is the error of uncertain parameter values caused by uncontrollable factors. Controllable factors, on the other hand, arise from human behavior. For instance, carelessness when gathering or processing data is considered to be a controllable factor. Some data utilized to estimate uncertain parameter values may thus be incorrect, implying that the estimated values will include errors. Uncertain parameter value error caused by controllable factors is defined as controllable error. Neither type of error can be ignored when implementing any numerical estimation or measurement experiments, for example, those carried out in general physics, according to Weltner et al. [31]. Uncontrollable error can be caused by interference factors during the experimental period, such as changes in air pressure, temperature or shocks. Controllable error on the other hand can be caused by discrepancies in the measurement instruments or approaches. Weltner et al. [31] also mentioned that both types of error have a notable effect on the accuracy of the measurement results. The uncertain parameter values in engineering optimization models may also contain both random and controllable errors. Therefore, the effects of these types of error in uncertain parameter values on the model solution must be evaluated to avoid making unsuitable decisions when solving such models.

An optimization model usually contains parameters in both the objective function and the constraints. In real world practices there may be uncertainty involved in both types of parameters. Recently, we developed an experimental method to evaluate solution errors from optimization models in which uncertain parameters were included in the objective function [32] and in the constraints [33], respectively. However, the results of error analysis for one type of uncertain parameters may differ from the results of error analysis for two types of uncertain parameters where a combined effect is usually incurred. Recently we thus conducted a prior study of error analysis for both types of uncertain parameters [34] and the test results preliminarily demonstrated the combined effect. Therefore, in this study, referring to Yan et al. [34] we develop an approach for a complete correctness evaluation, assuming uncertain parameters in both the objective function and the constraints for a project scheduling model. Different random and controllable error scenarios, with different solution tolerance error settings are used to understand the designs and the applications of optimization models and approximation solution algorithms. In cases where the model constraints contain errors due to uncertain parameter values there may be no feasible solution. The main reason for this being estimation errors in the parameter values for the constraints may alter the constraint set to become an empty set, resulting in infeasibility of the model. To obtain a feasible solution under this condition, the model has to be modified, as discussed in Section 4. In addition, the regression technique is adopted to construct equations which represent the relationships between input errors, solution tolerances and solution correctness (i.e., output errors).

To sum up, the main contributions of the study are:

(1) A method is proposed and an experiment performed to examine the influence of controllable and random errors leading to uncertain parameter values on the output of a project scheduling model where uncertain parameter values exist in both the objective function and the constraints. The method can also be applied to other optimization models, including maritime transport optimization models. A modified model is also developed to ensure model feasibility for performing all tests.
(2) Extensive tests are conducted to verify the effectiveness of the proposed method for models containing input errors in both the objective function and the constraints. In addition, the experiment is designed to include a number of error scenarios coupled with a number of solution tolerance errors.
(3) Regression analysis is adopted to create an equation for each error scenario to examine the relationship between the model input error, the solution tolerance error, and the model output error. Thus, decision makers can predict model output errors given model input errors and solution tolerance errors for similar models in practice.
(4) Some useful information and managerial suggestions based on the test results are proposed to assist decision makers in designing suitable strategies for dealing with similar problems in practice. For example, the design of employee training procedures or data handling processes to reduce model input errors or the setting of suitable solution tolerance errors so that model output errors and solution times can be reduced or controlled.

The rest of the paper is organized as follows. In Section 2, we introduce the project scheduling model. In Section 3, the modified model is described. In Section 4, we discuss uncertain parameters involved in the model. In Section 5, an approach to evaluate the model output error is proposed. In Section 6, error tests over uncertain parameter values and regression analysis of test results are carried out. Finally, some conclusions and suggestions for future work are given in Section 7.

2. Introduction to Chen et al.'s model

In this study we use the project scheduling model proposed by Chen et al. [21]; as an example, to perform the tests and to evaluate the effect of parameter uncertainties on the correctness of the results. There are two main reasons for choosing this model: first, it is easy to design various scenarios with controllable and random errors and to perform error tests over uncertain parameter values; second, extensive tests can be performed using 552 instances obtained from the project scheduling problem library (PSPLIB; http://www.om-db.wi.tum.de/psplib/main.html) associated with the model. Note that the proposed method can be applied to any optimization model which can be optimally solved using an exact solution algorithm. Studies on other models may be performed in the future. The model used here mainly deals with the multi-mode resource constrained project scheduling problem with discounted cash flows (MRCPSPDCF) with the PAC (payment at activity completion time) method for short term operations.

In this type of project scheduling problem, each project is made up of several activities. Each project must be concluded within an expected completion period. There are three factors that need to be considered when implementing the project: multiple modes, the use of renewable and non-renewable resources and the net present value (NPV). Activity preemption is not allowed. Each activity is carried out only by a single mode. Each mode has a specific duration and a specific consumption of renewable and non-renewable resources. The amount of renewable/non-renewable resources available for each time period/project period needs to be controlled. Each renewable/non-renewable resource has a specific cost. A discounted cash flow equation must be constructed to calculate the present value of the net cash flow of each activity. In Chen et al. [21]; all cash outflows for each activity are set at the activity starting time. All cash inflows for each activity conform to the PAC method. The aim of the MRCPSPDCF is to maximize the NPV of all cash flows for all activities in the project. Chen et al. [21] employed a time-precedence network flow technique to construct a generalized network flow model to optimally solve the MRCPSPDCF using the CPLEX11.1 mathematical programming software. For convenience, we briefly introduce the original model below. The interested reader may refer to Chen et al. [21] for a more detailed explanation.

Firstly, symbols and notations used in Chen et al.'s model are listed below.

Decision variable:

y_{ijk} : kth arc flow associated with node pair (i,j).

Parameters:

c_{ijk} : present value of net cash flow associated with arc (i,j,k);
r_{ijkl} : lth renewable resource amount associated with arc (i,j,k);
r_{ijko} : oth non-renewable resource amount associated with arc (i,j,k);
a_l : available lth renewable resource amount;
b_0 : available oth non-renewable resource amount;
s_{qi} : number of predecessors associated with node pair (q,i);
m_{qik} : flow adjustment coefficient associated with arc (q,i,k);
d : dth activity;
v : supply point;
f : collection point;
u_{ijk} : arc (i,j,k) flow upper bound.

Sets:

N : set of nodes;
D : set of activities;
W : set of node pairs associated with activities;
W_d : set of node pairs associated with the dth activity;
A_{ij} : set of parallel activity arcs associated with the node pair (i,j);
B_{qi} : set of arcs preceding the node pair (q,i);
RR : set of renewable resources;
NR : set of non-renewable resources;
T_h : set of node pairs associated with the hth time point;
T : set of time points associated with the project duration.
I : set of integers.

The model is shown as follows:

Maximize:

$$\sum_{ij\in W}\sum_{k\in A_{ij}} c_{ijk}y_{ijk} \tag{1}$$

Subject to:

$$\sum_{j\in N}\sum_{k\in A_{ij}} y_{ijk}-\sum_{q\in N}\sum_{k\in A_{qi}} m_{qik}y_{qik} = \begin{cases}1, & if \quad i=v\\ 0, & others \\ -1, & if \quad i=f\end{cases} \quad \forall i\in N \tag{2}$$

$$\sum_{k\in A_{ij}} s_{qi}y_{qik} \le \sum_{ij\in B_{qi}}\sum_{k\in A_{ij}} y_{ijk} \quad \forall (q,i)\in W \tag{3}$$

$$\sum_{ij\in W_d}\sum_{k\in A_{ij}} y_{ijk}=1 \quad d\in \mathrm{D} \tag{4}$$

$$\sum_{ij\in T_h}\sum_{k\in A_{ij}} r_{ijkl}y_{ijk}\le a_l \quad \forall h\in T, l\in RR \tag{5}$$

$$\sum_{ij\in W}\sum_{k\in A_{ij}} r_{ijko}y_{ijk}\le b_o \quad o\in NR \tag{6}$$

$$0\le y_{ijk}\le u_{ijk} \quad , \ x_{ijk}\in \ I \quad \forall k\in A_{ij}, \forall (i,j)\in W \tag{7}$$

Equation (1) is the objective function that maximizes the NPV of all cash flows for all activities in the project. Equation (2) ensures the flow conservation at every node in the network. Equation (3) denotes the precedence constraints between related activities. Equation (4) indicates that one mode only is selected for executing every activity. Equations (5) and (6) constrain the use of the available amount of renewable and non-renewable resources, respectively. Equation (7) ensures that all the arc flows are within their bounds.

Note that this model may be demonstrated to be infeasible when carrying out the error tests for the available amount of renewable and non-renewable resources. The model has to be suitably revised to alleviate this problem. The revised model is described in the next section.

3. The revised model

When carrying out the error tests for calculation of the available amount of renewable and non-renewable resources, the estimated available amount of resources could be less than the amount of resources required to finish the project (i.e., resource demand is greater than resource supply), so that the project cannot be completed (i.e., the model has no feasible solution). To avoid this occurring, Equations (1), (5) and (6) are revised to produce Equations (8), (9) and (12). The rest of the equations remain the same.

In Equations (8) and (9), extra amounts of the lth renewable resource and the oth non-renewable resource need to be added to the right-hand side in both equations, to complete the project. Notations δ_l and γ_o denote the extra amount of the lth renewable resource and the extra amount of the oth non-renewable resource required, respectively.

$$\sum_{ij\in T_h}\sum_{k\in A_{ij}} r_{ijkl} y_{ijk} \le a_l + \delta_l \quad \forall h \in T, l \in RR \tag{8}$$

$$\sum_{ij\in W}\sum_{k\in A_{ij}} r_{ijko}\, y_{ijk} \le b_o + \gamma_o \quad o \in NR \tag{9}$$

Two new Equations (10) and (11) are added to ensure the non-negativity and integrality of the extra amount of the lth renewable resource and the oth non-renewable resource, respectively.

$$\delta_l \ge 0 \& \delta_l \in Integer \quad l \in RR \tag{10}$$

$$\gamma_o \ge 0 \& \gamma_o \in Integer \quad o \in NR \tag{11}$$

In Equation (12), two penalty values, based on the extra amount of the lth renewable resource and the oth non-renewable resource required to complete the project, are added. Notations pv_l and pv_o represent the penalty value for extra use of a lth renewable resource and for extra use of a oth non-renewable resource, respectively.

$$Max \sum_{ij\in W}\sum_{k\in A_{ij}} c_{ijk} y_{ijk} - pv_l \delta_l - pv_o \gamma_o \tag{12}$$

4. Discussion of uncertain parameters included in the model

There is only one parameter (i.e., c_{ijk}) included in the objective function. Parameter c_{ijk} is used to express the NPV of all cash flows associated with each activity with a specific mode or the reward/penalty value associated with the collection arc. The NPV of all cash flows associated with each activity with a specific mode can be computed by a discounted cash flow equation comprised of three elements: the discount rate, all cash outflows for each activity with a specific mode and all cash inflows for each activity with a specific mode. The discount rate, a profit return value, is a certain parameter. In Chen et al. [21]; all cash outflows for each activity with a specific mode are equal to the mode cost. The mode cost can be calculated by multiplying the amount of renewable and non-renewable resources required by the mode by the unit cost of these resources. Since each mode is associated with a certain way of working, the amount of renewable and non-renewable resources required by each mode can be precisely evaluated, meaning that the amount of resources required by each mode will not contain errors. The unit cost of renewable and non-renewable resources usually requires the calculation of many complicated item costs, making it difficult to completely grasp and so it will contain errors (i.e., it is uncertain). In addition, all cash inflows for each activity with a specific mode are equal to the total contract payment for the project divided by the number of activities in the project. In practice, before executing the project, the decision maker can precisely evaluate the total contract payment and the number of activities in the project, meaning that they will not contain errors. The reward and penalty values for the collection arc will not contain errors, because they are mainly set by the decision maker.

There are six equations (i.e., Equations (2)–(7)) for the constraints. Equation (2) is used to ensure flow conservation for each node in the network and the values of parameter m_{qik} will not contain errors. Equation (3) is used to assure the precedence relationship for the activities and the values of parameter s_{qi} will not contain errors. Equation (4) is used to ensure that each activity is executed using a specific mode and the given parameter values will not contain errors. Equations (5) and (6) are used to manage the use of the renewable and non-renewable resources, respectively. The values of parameters a_l and b_o in these two equations may contain errors, because, in practice, to save time, only a rough estimate of the available amount of renewable and non-renewable resources required is made. In addition, the values of parameters r_{ijkl} and r_{ijko} will not contain errors, because each mode is associated with a way of working, meaning that the amount of renewable and non-renewable resources consumed by each mode can be precisely estimated. Equation (7) is used to define the bounds of all arc flows in the network, so the values of parameter u_{ijk} and the given parameter values will not contain errors.

To conclude the above analysis, the unit cost as well as the available amounts of renewable and non-renewable resources are uncertain parameters. Therefore, in this study, the error tests are mainly carried out over these uncertain parameter values.

5. An approach for evaluating model output errors

In this section an approach for evaluating the solution correctness is proposed in order to explore the output errors of Chen et al.'s model which contains uncertain parameters in both the objective function and the constraints. As discussed in Section 4, the exploration of model output errors carried out in the approach is mainly on the revised model. Note that because many of the parameters which could exist in an optimization model with controllable and random errors, as in Chen et al.'s model, lack a specific pattern, traditional sensitivity analysis and parametric programming techniques cannot be used to examine the effect of the parameter errors on the model output. Generally, when performing the evaluation of uncertain parameter values, the existence of both controllable and random errors has to be considered. Hence, several different input error scenarios are designed to represent different conditions.

Each error scenario is associated with a specific error range which is adjustable. The error range is used to generate controllable and random errors. For instance, when the error range is set to ±3%, the generated controllable and random errors will be in the range of ±3%. According to Kline and McClintock [35] the generation of controllable and random errors conforms to a normal distribution. In this approach, the generation of controllable and random errors associated with all designed error scenarios mainly conforms to a truncated standard normal distribution (the two tails for the distribution are separately cut by 5%) in order to avoid generating extreme errors. If only one set of errors is produced for each scenario, the produced error may be excessively subjective. To increase the objectivity, each error scenario will be designed to produce multiple sets of errors. The evaluation of a suitable set of errors for each error scenario will be discussed in Section 6.2. In addition, we can use each set of errors produced in each error scenario coupled with the real value of uncertain parameters to calculate the corresponding value of uncertain parameters for the error scenario. A simple example is used to explain how to use the set of errors produced in an error scenario coupled with the real value of an uncertain parameter to calculate the corresponding value of the uncertain parameter. In this example, the value of an uncertain parameter e needs to be evaluated. The real value of uncertain parameter e is assumed to be 200. There is only one set of controllable and random errors generated in the error scenario. The error range for the error scenario is set to be ±3%. The generated controllable and random errors are 0.02 and −0.01, respectively. In the error scenario, the value of e can be obtained by taking the real value of uncertain parameter e plus the value influenced by the generated controllable and random errors. Therefore, the value of e in the error scenario is 202 (= 200 + 200*(0.02 - 0.01)). There is a model solution corresponding to each set of errors generated in each error scenario. Since there are multiple sets of errors generated for each error scenario, each error scenario has multiple corresponding model solutions. In this approach, multiple model solutions for each error scenario must be averaged to become a representative model solution associated with the error scenario.

The effect of the solution tolerance errors on the model output errors is also examined. Various solution tolerance errors are designed when solving the revised model with uncertain parameter values for each error scenario. An equation is proposed to compute the gap between the real optimal solution and the solution obtained from the revised model with uncertain parameter values for the specific error scenario and the specific solution tolerance error setting, as discussed below. In the equation, for convenience, the solution obtained from Chen et al.'s model with real parameter values, and a solution tolerance error of 0% is regarded as the real optimal solution. Parameter value settings for the revised model will be introduced in Section 6.1. The real parameter value settings are the same as those set in Chen et al. [21]. In the equation, the notation RS indicates the representative model solution for the revised model with uncertain parameter values under the specific error scenario, coupled with the specific solution tolerance error setting and OS indicates the real optimal solution.

$$\text{Model output errors}(\%) = \left|\frac{RS - OS}{OS}\right| \times 100\%$$

6. Error tests for uncertain parameter values

Error tests for uncertain parameter values are carried out using 552 test instances for a 30-activity project scheduling problem, as used in Chen et al. [21]. The data for the test instances can be obtained directly from the project scheduling problem library (http://www.om-db.wi.tum.de/psplib/main.html). The Visual C++ programming language, coupled with CPLEX 11.1 (the solution procedure is comprised of the branch and bound method, coupled with the simplex method) is used to construct and solve the model. A Celeron-M540 2.30 GHz CPU with 2.0 GB of RAM operating in the environment of Microsoft Windows XP is utilized to perform the tests.

6.1. Parameter value settings for the revised model

The same parameter value settings are used for all test instances in this study. To save space, a test instance with a PSPLIB file number of mf11_.bas is utilized to introduce the parameter value settings for the revised model. In the test instance, three modes are provided for executing each activity. The time period of each mode is between 1 and 10 time units. The time units are adjustable (e.g., it can be set to be a quarter of a year, half a year, or a year). The type of renewable resource is two (i.e., R1 and R2). The type of non-renewable resource is two (i.e., N1 and N2). The expected finish period of the project is 35 time units. The available amounts for R1/R2 per time unit are 32/28. The available amounts for N1/N2 in the analysis period are 86/93. The analysis period (i.e., the longest execution period of the project) is 78 time units, which is computed by the label-correcting algorithm used in Chen et al. [21]. For the period and the resource demand related to the mode, as well as the successors for each activity, please refer to the file. There are some data (e.g., the cost parameter data for the activities, the permissible work periods for the activities, and the data for the dummy activity arcs) that are not provided in the file. For convenience, these data are mainly identical to those used by Chen et al. [21]. The contract payment for the project is 4,000,000 monetary units. The monetary unit is adjustable (e.g., it can be set to be USD or NTD). All cash inflows for each activity with all modes are equal to the contract payment for the project divided by the number of activities in the project. All cash inflows for each activity with all modes in the test instance can be computed to be about 133,333 (= 4,000,000/30) monetary units. All cash outflows for each activity are equal to the cost of the mode performing the activity. The collection arc cost is 4000 monetary units per time unit. The discount rate is 0.03 per time unit. In addition, we employ the methods proposed in Chen et al. [21] to build the dummy activity arcs and calculate the permissible work periods for the activities.

In accordance with the analysis in Section 3, in Chen et al.'s model, the unit cost and the available amounts for R1/R2/N1/N2 are considered to be uncertain parameters. Accordingly, we design 15 error scenarios for the unit cost and available amounts for R1/R2/N1/N2; see Table 1. No error exists for the uncertain parameter values in error scenario 1. Since the change in the controllable error range is more obvious than in the random error range for a short-term planning model, there are two random error ranges (i.e., 0% and ±5%) set and seven controllable error ranges (i.e., ±1%, ±3%, ±5%, ±7%, ±10%, ±15% and ±20%) set in the other error scenarios. Note that the unit costs for resources should not be significantly variated within the short period in practice. Consequently, the range of random errors for these costs is expectedly small, compared with that of the controllable errors. That is why the maximum random error is set smaller than the maximum controllable error. Note that the ranges of random errors and controllable errors, which are associated with the problems and the environment, are adjustable in real practices. This study does not evaluate the influence of purely random errors on model output errors, because we want to focus on understanding the influence of controllable errors on output errors. In the future, the influence of purely random errors can be evaluated. The real value of the unit cost for R1/R2/N1/N2 is set to be 13,065/5823/11,628/5538 monetary units, the same as that set in Chen et al. [21]. The real value of the available amount for R1/R2/N1/N2 is set to be the same as that given for the test instance. The penalty value used for the extra use of R1/R2 is set to be 1.5 times the unit cost for R1/R2 (i.e., 19,598/8735 monetary units). The penalty value used for the extra use of N1/N2 is set to be 2 times

Table 1. Error scenarios for the uncertain parameter values.

Error scenario	Controllable error range	Random error range
1	–	–
2	±1%	–
3	±3%	–
4	±5%	–
5	±7%	–
6	±10%	–
7	±15%	–
8	±20%	–
9	±1%	±5%
10	±3%	±5%
11	±5%	±5%
12	±7%	±5%
13	±10%	±5%
14	±15%	±5%
15	±20%	±5%

Note: The notation "–" denotes that no error range exists.

the unit cost for N1/N2 (i.e., 23,256/11,076 monetary units). The main reason for this is that renewable resources are easier to obtain than non-renewable resources.

The test instance contains a time-precedence network. There are 79 time-precedence points associated with each precedence point in the analysis period. The revised model includes 2028 nodes, 6151 arcs and 14,511 constraints, in which 2028 constraints are used to ensure the flow conservation of all nodes, 6151 constraints are used to ensure the all arc flow bounds, and 6332 side constraints are used to comply with the operating regulations. In terms of solving the revised model under each error scenario, we design 11 solution tolerance errors, from 0% to 10%, in increments of 1%.

6.2. Output results

To ascertain how many error sets need to be generated for each of the error scenarios ranging from 2 to 15, the evaluation of a proper error set must be carried out. For convenience, error scenario 11 is evaluated, with zero solution tolerance error. Ten situations, from 40 to 130 scenarios, increasing in increments of 10 scenarios, are tested in the evaluation. Each scenario has a specific set of random and controllable errors. The test instance used here is the same as that explained in Section 6.1. As shown in Fig. 1, after 100 scenarios, the average objective values stabilize. For ease of testing, when solving each test instance for each problem, with all solution tolerance error settings, 100 error scenarios are generated for each of error scenarios 2 to 15.

Table 2 shows the output results for a project that includes 30 activities under different error scenarios, with different solution tolerance error settings. For error scenario 1 in Table 2, we show the average real solution errors (%) and average solution time

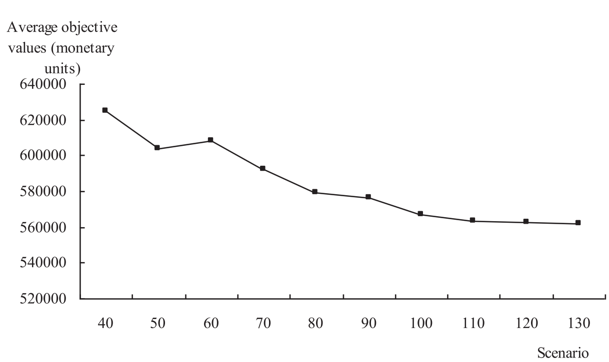

Fig. 1. Average objective values for various numbers of scenarios.

Table 2. Output results for each error scenario.

Error scenario	Solution tolerance error (%)										
No.	0	1	2	3	4	5	6	7	8	9	10
1	0 (31.27)	0.41 (27.79)	0.75 (24.09)	1.54 (20.73)	2.33 (19.60)	2.75 (18.85)	3.21 (18.11)	3.79 (17.87)	4.62 (17.31)	5.23 (16.73)	5.97 (15.42)
2	0.74 (32.10)	0.88 (27.93)	0.98 (24.83)	1.09 (21.63)	1.62 (19.82)	1.84 (18.50)	2.71 (18.05)	3.26 (17.37)	3.81 (16.53)	4.34 (16.17)	5.65 (14.99)
3	2.18 (33.26)	2.32 (28.30)	2.42 (26.14)	2.91 (23.85)	3.13 (21.01)	3.42 (19.71)	3.92 (18.04)	4.44 (16.89)	4.88 (16.43)	5.75 (15.29)	6.99 (14.46)
4	3.57 (32.17)	3.80 (27.74)	3.92 (24.34)	4.14 (22.50)	4.62 (20.18)	4.94 (18.52)	5.55 (17.85)	6.09 (17.32)	6.61 (16.97)	7.35 (16.61)	8.74 (16.01)
5	5.36 (31.79)	5.45 (28.72)	5.62 (25.47)	5.91 (22.45)	6.53 (20.21)	7.11 (19.18)	7.83 (18.10)	8.14 (17.83)	8.63 (17.37)	9.56 (17.14)	10.62 (16.16)
6	6.94 (33.11)	7.21 (29.67)	7.39 (26.72)	7.62 (22.56)	8.64 (19.91)	9.17 (19.08)	10.07 (18.71)	10.95 (18.49)	11.53 (18.18)	12.61 (17.79)	14.44 (16.92)
7	11.32 (32.82)	11.41 (29.22)	11.50 (24.94)	12.03 (23.03)	12.92 (20.85)	13.99 (19.79)	14.87 (19.14)	15.46 (18.66)	16.24 (18.13)	17.45 (16.73)	18.57 (16.28)
8	17.43 (31.47)	17.52 (27.59)	17.62 (24.36)	17.85 (22.16)	18.61 (19.07)	19.16 (18.35)	19.75 (17.70)	20.26 (17.53)	21.09 (16.78)	21.84 (16.33)	23.53 (15.67)
9	3.23 (31.59)	3.37 (28.07)	3.51 (24.31)	3.72 (20.94)	3.99 (19.80)	4.06 (19.04)	4.47 (18.30)	5.09 (18.05)	5.38 (17.48)	5.51 (16.90)	6.31 (15.57)
10	4.01 (33.26)	4.11 (30.59)	4.33 (27.19)	4.67 (24.22)	4.95 (20.55)	5.74 (18.75)	5.93 (17.02)	6.14 (16.19)	6.54 (15.53)	6.96 (14.93)	7.81 (14.46)
11	7.28 (32.93)	7.59 (28.41)	7.85 (26.29)	8.12 (22.02)	8.93 (20.53)	9.82 (19.37)	10.35 (18.47)	10.68 (17.31)	11.25 (16.79)	11.54 (16.58)	12.55 (15.64)
12	10.41 (32.24)	10.65 (28.95)	10.81 (26.76)	11.14 (22.29)	12.12 (20.48)	12.31 (19.57)	12.63 (17.19)	13.45 (16.96)	14.36 (16.76)	15.46 (16.70)	16.17 (16.63)
13	12.36 (31.76)	12.44 (28.38)	12.73 (25.01)	13.04 (22.55)	14.03 (20.29)	14.85 (18.85)	15.58 (18.40)	16.02 (17.88)	16.56 (17.01)	18.25 (16.54)	20.62 (16.40)
14	14.72 (32.26)	14.98 (28.50)	15.48 (24.09)	15.83 (21.53)	16.84 (19.84)	17.65 (19.10)	18.34 (18.28)	19.03 (17.94)	19.95 (17.69)	20.65 (17.25)	23.04 (16.21)
15	20.11 (32.39)	20.32 (28.58)	20.58 (25.25)	20.75 (22.44)	21.83 (21.01)	22.42 (19.91)	23.25 (19.22)	23.97 (18.13)	24.78 (17.23)	26.07 (15.90)	28.09 (14.34)

Note: The number outside the brackets associated with an error scenario and a solution tolerance error indicates the average real solution error (also the average model output error) (%) for all test instances; the number inside the brackets associated with an error scenario and a solution tolerance error denotes the average solution time (seconds) for all test instances.

(seconds) for all test instances under each solution tolerance error setting. Since no errors exist in uncertain parameter values for error scenario 1, the average real solution error is zero when the solution tolerance error is set to be 0%, meaning that the solutions are real optimal solutions. In addition, the larger the solution tolerance error set, the larger the average real solution error and the shorter the average solution time. In contrast, the smaller the solution tolerance error set, the smaller the average model output error and the longer the average solution time.

Figs. 2–5 show more detailed results for error scenarios 2 to 15. Note that in these four figures, the presented solution tolerance errors are the solution tolerance error settings in CPLEX 11.1. Fig. 2 and 3 show that the model output errors for each of error scenarios 2 to 15 are still positive when the solution tolerance error is set to 0%. Based on this, we find that when the engineering optimization model contains uncertain parameters (i.e., model input contains errors), the model solution is not optimal (i.e., model output contains errors), even though the solution tolerance error is set to 0%. This means that any decision made using the solution obtained from an engineering optimization model with uncertain parameters is not optimal. On the other hand, the lower the error range set, the smaller the average model output error shown for error scenarios 2 to 8/9 to 15, with a solution tolerance error setting of 0%. Based on this, we find that the model output error can be expected to be lowered by increasing the precision of uncertain parameter values.

Fig. 4 shows the relative increase in average model output errors regarding all error scenarios 2–15 as the solution tolerance error increases. It is found that when the solution tolerance error is less than 3%, the relative increase in average model output errors is not significant. Fig. 5 shows the relative decrease in average solution time for all error scenarios 2–15 as the solution tolerance error increases. It is found that under a solution tolerance error of 3%, the relative decrease in average solution time for these error scenarios is significant. The implication is that when solving Chen et al.'s model or similar models, with uncertain parameter values, with approximation solution algorithms, the solution tolerance error should be set to about 3%, to maintain solution correctness while significantly shortening the solution time. Note that suitable solutions tolerance error settings may differ depending upon whether CPLEX or other solution algorithms are used to solve other models with uncertain parameters and further testing is needed. T-test is carried out to understand whether the model output errors for each of error scenarios 2–15 with 0% solution tolerance error are greater than those obtained by [32,33]; where only uncertain parameters are included in either the objective function or the constraints. The significance level α is set to be 1%. The results indicate that the p-values of all t-tests are all below the significance level α of 1%. This shows that the model output errors for each of error scenarios 2–15 obtained in this study are greater than those obtained in Yan et al. [32]/Yan et al. [32]. As a result, when more uncertain parameters are included in the model, there will be more model output errors generated.

6.3. Regression analysis

We use the SAS 9.3 software to perform regression analysis of the test results for each error scenario in order to understand the relationship between model input error, solution tolerance error and model output error. There are two independent

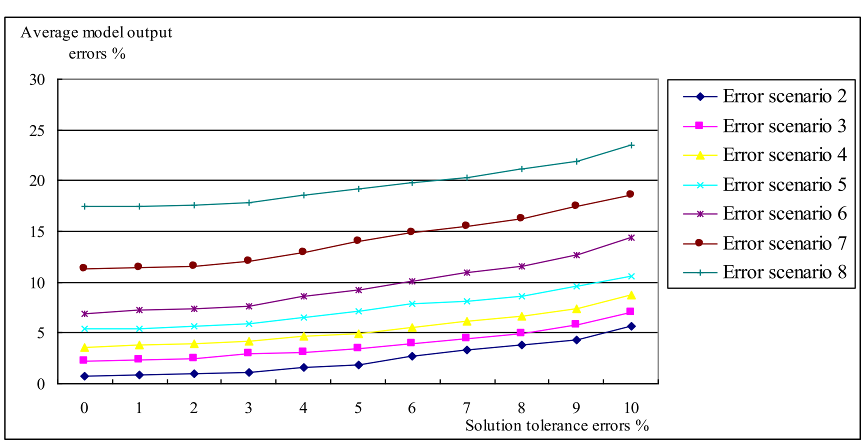

Fig. 2. Average model output errors for each of error scenarios 2–8 with various solution tolerance errors.

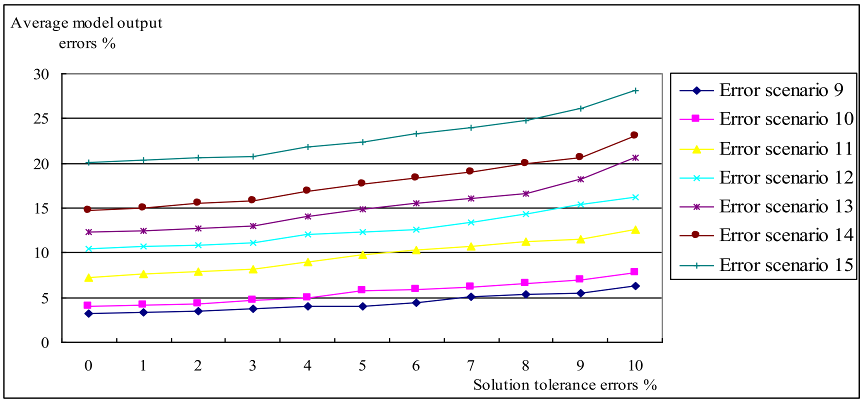

Fig. 3. Average model output errors for each of error scenarios 9−15 with various solution tolerance errors.

variables x_1 and x_2 (i.e., model input error and solution tolerance error) and one dependent variable y (i.e., model output error) in the regression analysis. A test of the independence associated with x_1 and x_2 will be carried out later. The regression equation is estimated with respect to an error scenario. The model output error can be predicted given the controllable input error and the solution tolerance error associated with the regression equation. Since there is no model input error in error scenario 1, there is only one independent variable x_2 in the corresponding regression equation. The regression equations corresponding to each of error scenarios 2 to 15 include two independent variables x_1 and x_2. To assure the appropriateness of the regression equation associated for each error scenario, the powers for x_1 and x_2 need to be evaluated. The powers of x_1 and x_2 are first set to be 1. We found most of the R-square values associated with the equations to be less than 0.3, meaning that they were not good enough to predict model output errors. We then proposed a simple way of evaluating the proper powers for x_1 and x_2 as follows. First, error scenarios 2 to 15 are divided into two categories. One contains error scenarios 2 to 8 where x_1 includes controllable error only. The other contains error scenarios 9 to 15 where x_1 includes both random and controllable errors. We carry out t-test on the test results of the two categories, where the significance level α is set to be 1%, to understand whether the test results for the two categories will be different. The results indicate that the p-values of the t-test are below the significance level α of 1%, showing that there are notable differences in the test results between error scenarios 2−8 and error scenarios 9−15. Now, regression analysis of all test results of error scenarios 2 to 8/9 to 15 for each problem is carried out. Various powers for x_1 and x_2 are tested, from 0.6 to 1.5, in increments of 0.1. The intercept of the regression equation is set to zero. After performing the tests, we choose the powers for x_1 and x_2 so that the corresponding regression equation has the best adjusted R-square value for each of error scenarios 2−8/9-15. The power for x_2

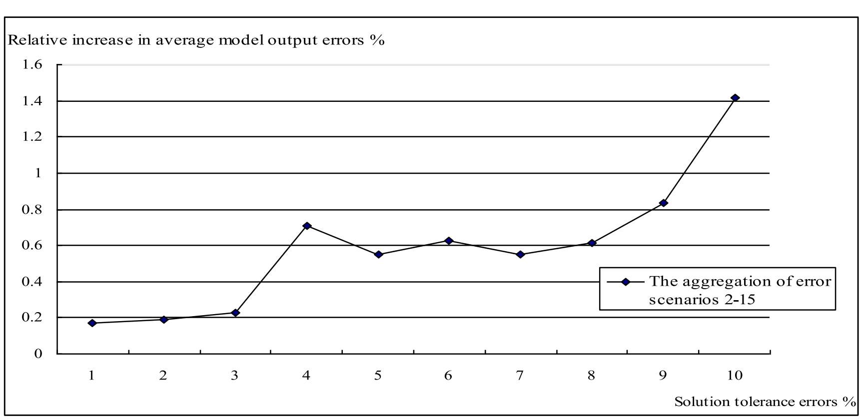

Fig. 4. Relative increase in average model output errors for error scenarios 2−15 as solution tolerance error increases.

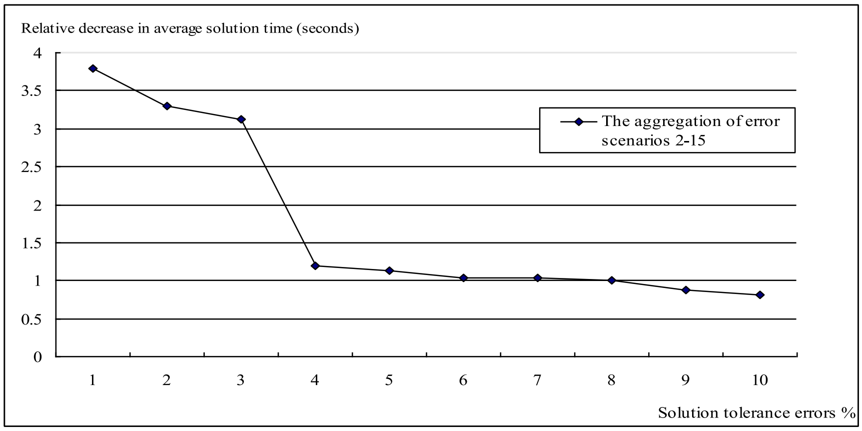

Fig. 5. Relative decrease in average solution time for error scenarios 2–15 as solution tolerance error increases.

for the regression equation for error scenario 1 is set to be the same as that for the regression equation with the best adjusted R-square value for error scenarios 2 to 8.

The estimated regression equations and the corresponding adjusted R-square values for all error scenarios are presented in Table 3. For each equation, the sample size is 6072 (= 552*11; the first number indicates the number of test instances and the latter the number of solution tolerance errors). According to the proposed method, the best adjusted R-square value for the estimated regression equation for error scenarios 2 to 8/9 to 15 is obtained when the powers for x_1 and x_2 are equal to 1.3/1.4 and 0.8/0.8. The controllable error is amplified more than the solution tolerance error under a random error setting. Note that the power for x_2 in the estimated regression equation associated with error scenario 1 is set to be 0.8. Based on the Pearson correlation coefficient test, the two independent variables in the estimated regression equation for each of error scenarios 2 to 15 are independent. The results of t-test with a significance level (i.e., α) of 1% confirm the parameter estimates of the estimated regression equation for each error scenario to be reliable. The results of F-test with a significance level (i.e., α) of 1% confirm the estimated regression equation for each error scenario to be reliable. In addition, the adjusted R-square value for the estimated regression equation for each error scenario is good (>0.8) and acceptable.

The two independent variables (i.e., x_1 and x_2) in the estimated regression equation for each of error scenarios 2 to 15 positively affect the dependent variable (i.e., y). The effect of x_1 on y slightly decreases for most estimated regression equations for error scenarios 2 to 8/9 to 15. It can also be found that the coefficient values of x_1 for the equations corresponding to error scenarios 2 to 8 are greater than those for the equations corresponding to error scenarios 9 to 15, since with larger random errors, the effect of solution tolerance error on model output is diluted. Besides, the coefficient values of x_2 for most equations corresponding to error scenarios 9 to 15 are slightly less than those for the equations corresponding to error scenarios 2 to 8, mainly

Table 3. Estimated regression equation for each error scenario.

Error scenario No.	Estimated regression equation	Adjusted R-square	Error scenario No.	Estimated regression equation	Adjusted R-square
1	$y = 1.28x_2^{0.8} + \epsilon$	0.8635	9	$y = 0.41x_1^{1.4} + 0.89x_2^{0.8} + \epsilon$	0.8831
2	$y = 0.87x_1^{1.3} + 0.83x_2^{0.8} + \epsilon$	0.9136	10	$y = 0.34x_1^{1.4} + 1.04x_2^{0.8} + \epsilon$	0.8991
3	$y = 0.72x_1^{1.3} + 0.96x_2^{0.8} + \epsilon$	0.9054	11	$y = 0.45x_1^{1.4} + 1.51x_2^{0.8} + \epsilon$	0.8725
4	$y = 0.65x_1^{1.3} + 1.45x_2^{0.8} + \epsilon$	0.8974	12	$y = 0.50x_1^{1.4} + 1.62x_2^{0.8} + \epsilon$	0.9036
5	$y = 0.63x_1^{1.3} + 1.61x_2^{0.8} + \epsilon$	0.9131	13	$y = 0.44x_1^{1.4} + 2.16x_2^{0.8} + \epsilon$	0.9012
6	$y = 0.56x_1^{1.3} + 2.07x_2^{0.8} + \epsilon$	0.9058	14	$y = 0.39x_1^{1.4} + 2.31x_2^{0.8} + \epsilon$	0.8957
7	$y = 0.52x_1^{1.3} + 2.21x_2^{0.8} + \epsilon$	0.9022	15	$y = 0.42x_1^{1.4} + 2.69x_2^{0.8} + \epsilon$	0.8818
8	$y = 0.61x_1^{1.3} + 2.53x_2^{0.8} + \epsilon$	0.8938	9–15	$y = 0.38x_1^{1.4} + 1.77x_2^{0.8} + \epsilon$	0.8521
2–8	$y = 0.60x_1^{1.3} + 1.87x_2^{0.8} + \epsilon$	0.8729			

Note: The symbol $y/x_1/x_2/\epsilon$ represents model output error/model input error/solution tolerance error/error item. For each equation, the p-values from t-test are below the significance level α of 1%, indicating that the parameter estimated for each equation is significant, and the p-values from F-test are below the significance level α of 1%, indicating that each equation is significant.

because the coefficient values of x_2 for the equations corresponding to error scenarios 9 to 15 are slightly influenced by the random errors. For the equations corresponding to error scenarios 2 to 8/9 to 15, the coefficient values of x_2 are greater than those of x_1, but the power of x_2 is less than that of x_1. With the combined power effect, the influence of x_1 on y will be greater than that of x_2 on y. In addition, the powers of x_1 for the estimated equations for error scenarios 9 to 15 are greater than those for error scenarios 2 to 8, indicating that an environment with higher random errors could amplify model output errors. The effect of x_2 on y increases for the estimated equations for error scenarios 2 to 8/9 to 15 when the controllable error range associated with uncertain parameter values increases. This implies that, in an environment with higher random errors, the setting of solution tolerance errors must be carefully determined to avoid the occurrence of excessive model output errors. Finally, these regression equations could be useful for decision makers to predict possible model output errors given model input errors and solution tolerance errors for similar models.

6.4. Findings and discussion of the output results

Some important findings obtained from the output results are detailed below.

(1) When an engineering optimization model includes uncertain parameters, the model output will contain errors even though the solution tolerance error is set to 0%. Decisions made with this solution will not be optimal, although it has always been thought that they were optimal. Moreover, for solving engineering optimization models with uncertain parameters, the fewer the model input errors set, the smaller the model output errors generated. In other words, we can expect to decrease output errors by reducing model input errors by enhancing the accuracy of uncertain parameter values. Thus, decision makers should design data handling processes or train employees to reduce model input errors. This would reduce model output errors.

(2) In order to shorten the solution time while maintaining solution correctness, the solution tolerance error should be set to 3% or 4% for Chen et al.'s or similar models that contain uncertain parameters to solve real problems in practice. This ensures that the obtained objective values will not be significantly different from the exact ones, while the solution time could be saved by 30%. Note that the most appropriate solution tolerance error settings for other optimization models with uncertain parameters may be different and can be similarly examined in the future. Note also that it is usually necessary to utilize approximation solution algorithms with a solution tolerance error to efficiently solve engineering optimization models that are characterized as NP-hard [36]. Decision makers may also decide how to set the solution tolerance errors to solve for near-optimal solutions to control model output errors.

(3) Many reliable regression equations are estimated using the test results obtained in this study from which many insights can be obtained. For example, when the given input error range increases, there is a slight decrease in the effect of the model input error on the output errors for most regression equations for error scenarios 2 to 8/9 to 15 (i.e., model input errors will not significantly affect model output errors). The coefficient values of the model input error for the equations corresponding to error scenarios 2 to 8 are greater than those for the equations corresponding to error scenarios 9 to 15, because random errors dilute the effect. For another example, for the equations corresponding to error scenarios 2 to 8/9 to 15, the coefficient values of the solution tolerance error are greater than those for the model input error, but the power of the solution tolerance error is less than that of the model input error. The combination effect makes the effect of model input error greater than that of solution tolerance error on model output error. Decision makers can choose the most appropriate equation for the prediction of possible model output errors when solving the engineering optimization models that are similar to Chen et al.'s model with uncertain parameters.

(4) For the regression equations for error scenarios 2 to 8/9 to 15, when the error range associated with uncertain parameter values increases, the effect of solution tolerance errors on model output errors increases. This implies that if random errors are considered, then the setting of solution tolerance errors must be carefully determined in order to avoid the occurrence of excessive model output errors. In addition, the powers of model input errors for the equations for error scenarios 2 to 8 are less than those for error scenarios 9 to 15, indicating that a decrease in model input errors will lower the amplified effect on model output errors. For the equations for error scenarios 2 to 8/9 to 15, a decrease in the number of uncertain parameters will cause

that the powers of the model input errors drop. Thus, when solving engineering optimization models that are similar to Chen et al.'s model with uncertain parameters, model input errors and the number of uncertain parameters should be lowered as much as possible in order to lessen the amplified effect on model output errors.

(5) The patterns of relationship between the input error, solution tolerance and output error are similar to those obtained by [32,33]; who focused on parameter uncertainty in the objective function and in the constraints, respectively, verifying the usefulness of the proposed method. However, after carrying out t-test of model output errors for each of error scenarios 2–15 and making a comparison between the results obtained in this study and in Yan et al. [32]/Yan et al. [33]; we find that the p-values of all t-tests are below the significance level α of 1%. This shows that the model output errors for error scenarios 2–15 obtained in this study are greater than those obtained in Yan et al. [32]/Yan et al. [33]; meaning that when more uncertain parameters are included in the model, there will be more model output errors generated, mostly due to the combined effect of parameter uncertainty in the objective function and in the constraints.

7. Conclusions

This study develops an approach, using Chen et al.'s [21] model as a testbed, to examine the solution correctness of optimization models with uncertain parameter values under various controllable and random error scenarios in the objective function and in the constraints, with various solution tolerance error settings. To model is revised to ensure feasibility in all tests. To verify the proposed approach, we design 15 error scenarios for controllable and random errors, and 11 solution tolerance errors. With 552 test instances, there are a total of 91,080 (=15*11*552) tests in the experiment. Regression analysis of the output results is carried out in order to understand the relationship between the model input error, the solution tolerance error and the model output error. Useful findings are obtained which should help decision makers adopting Chen et al.'s model or similar models in real world practices. For example, to shorten the solution time while maintaining solution correctness, the solution tolerance error may be set as 3% or 4% when applying models with uncertain parameters in practice. However, the solution tolerance settings may be different for other optimization models with uncertain parameters and can be similarly examined in the future.

Many reliable regression equations are estimated using the test results obtained in this study from which many insights can be obtained. For example, random errors dilute the effect of model input errors on model output errors, but the effect is still greater than that of solution tolerance error on model output error. For another example, the solution tolerance errors should be carefully set to avoid the occurrence of excessive model output errors when there are additional random errors contained in the model input errors. Moreover, the combined effect of parameter uncertainty in the objective function and in the constraints on the solution correctness is higher than the individual effect of parameter uncertainty in either the objective function or the constraints. Decision makers can select the most appropriate equation for predicting possible model output errors, and adopt strategies to reduce the model output error, when solving the engineering optimization models that are similar to Chen et al.'s model with uncertain parameters. Finally, although the obtained results and findings in this study could be applied to Chen et al.'s model or similar models, the proposed application process could also be useful for the decision makers to examine other optimization models, including maritime transport optimization models, in the future, so that these models can be more efficiently and effectively applied in practice.

Acknowledgments

This research was supported by two grants (MOST-104-2221-E-008-032-MY3 and MOST-107-2221-E-008-018-MY3) from the Ministry of Science and Technology of Taiwan. The authors would like to thank two anonymous reviewers for their helpful comments and suggestions on the presentation of the paper.

References

[1] An N, Zhao W, Wang J, Shang D, Zhao E. Using multi-output feedforward neural network with empirical mode decomposition based signal filtering for electricity demand forecasting. Energy 2013;49:279–88.

[2] Chen CF, Lai MC, Yeh CC. Forecasting tourism demand based on empirical mode decomposition and neural network. Knowl Base Syst 2012;26:281–7.

[3] Chiang WY. Establishment and application of fuzzy decision rules: an empirical case of the air passenger market in taiwan. Int J Tourism Res 2011;13:447–56.

[4] Choi SY. Short-term power demand forecasting using information Technology based data mining method. Comput Sci ITS Appl 2006;3984:322–30.

[5] David JS, Riham AK, Lawrence MM. Cost model development using virtual manufacturing and data mining: Part

II—comparison of data mining algorithms. Int J Adv Manuf Technol 2013;66:1389−96.
[6] Gholamreza Z, Saeed A, Alireza B, Ali E, Sharifah RWA. Electricity demand estimation using an adaptive neuro-fuzzy network: a case study from the ontario province-Canada. Energy 2013;49:323−8.
[7] Hsu CI, Wen YH. Application of grey theory and multi-objective programming towards airline network design. Eur J Oper Res 2000;127:44−68.
[8] Huang YL, Lee YH. Accurately forecasting model for the stochastic volatility data in tourism demand. Mod Econ 2011; 2:823−9.
[9] Jin JG, Nieto H, Lu L. Robust bike-sharing stations allocation and path network design: a two-stage stochastic programming model. Transport Lett 2020;12(10):682−91.
[10] Kenyon AS, Morton DP. Stochastic vehicle routing with random travel times. Transport Sci 2003;37:69−82.
[11] Lee WH, Tseng SS, Shieh JL, Chen HH. Discovering traffic bottlenecks in an urban network by spatiotemporal data mining on location-based services. IEEE Trans Intell Transport Syst 2011;12:1047−56.
[12] Lin FT, Yao JS. Using fuzzy numbers in knapsack problems. Eur J Oper Res 2001;135:158−76.
[13] Lu CC. Robust multi-period fleet allocation models for bike-sharing systems. Netw Spatial Econom 2016;16:61−82.
[14] Parnianifard A, Azfanizam A, Ariffin M, Ismail M. An overview on robust design hybrid metamodeling: advanced methodology in process optimization under uncertainty. Int J Ind Eng Comput 2018;9(1):1−32.
[15] Yan S, Lu CC, Hsieh JH, Lin HC. A dynamic and flexible berth allocation model with stochastic vessel arrival times. Network Spatial Econ 2019;19(3):903−27.
[16] Yan S, Wang SS, Chang YH. Cash transportation vehicle routing and scheduling under stochastic travel times. Eng Optim 2014;46(3):289−307.
[17] Zhao H, Zheng J, Xu J, Wu D. Fault diagnosis method based on principal component analysis and broad learning system. IEEE Access 2019;7:99263−72.
[18] Opgenoord MMJ, Allaire DL, Willcox KE. Variance-based sensitivity analysis to support simulation-based design under uncertainty. J Mech Des 2016;138(11):111410−21.
[19] Li L, Lu Z. Variance-based sensitivity analysis for models with correlated inputs and its state dependent parameter solution. Struct Multidiscip Optim 2017;56(4):919−37.
[20] Hillier GJ, Lieberman FS. Introduction to mathematical programming. New York City, NY: McGraw-Hil; 1995.
[21] Chen M, Yan S, Wang SS, Liu CL. A generalized network flow model for the multi-mode resource constrained project scheduling problem with discounted cash flows. Eng Optim 2015;47(2):165−83.
[22] Chen D, Wu H. Research on optimization model and algorithm of initial schedule of intercity passenger trains based on fuzzy sets. J Softw 2012;7:49−54.
[23] Chu CW, Chang JR, Chang KH. A heuristic algorithm for the multi-depot vehicle routing problem with outsider carrier selection. J Marine Sci - Taiwan 2019;27:170−83.
[24] Parpinelli RS, Lopes HS, Freitas AA. Data mining with an ant colony optimization algorithm. IEEE Trans Evol Comput 2002;6(4):321−32.
[25] Wang CH. Predicting tourism demand using fuzzy time series and hybrid grey theory. Tourism Manag 2004;25:367−74.
[26] Wanigasooriya J, Gifernando T. Multi-vehicle passenger allocation and route optimization for employee transportation using genetic algorithms. Int J Comput Appl 2013; 64:1−9.
[27] William CC, Asunción PC, Antonio FC. Optimal design of energy-efficient ATO CBTC driving for metro lines based on NSGA-II with fuzzy parameters. Eng Appl Artif Intell 2014; 36:164−77.
[28] Wu CS, Chu CW, Hsu HL. A heuristic algorithm of vehicle routing problem with time Windows and less-than-truckload carrier selection. J Marine Sci - Taiwan 2017;25(2):129−41.
[29] Yan S, Tang CH, Fu TC. An airline scheduling model and solution algorithms under stochastic demands. Eur J Oper Res 2008;190:22−39.
[30] Zhu C, Li H. Research on optimization algorithm based on random gray ant colony neural network and its applications. J Inf Comput Sci 2012;9:2475−83.
[31] Weltner K, Weber WJ, Grosjean J, Schuster P. Theory of errors. Math Phys Eng 2009:537−56.
[32] Yan S, Chu JC, Wang SS. An experimental approach for examining solution errors of engineering problems with uncertain parameters. Comput Ind Eng 2017;113:1−9.
[33] Yan S, Wang SS, Wang CY. The influence of constraint uncertainty on model solution correctness. J Chin Inst Eng 2021; 44(4):319−31.
[34] Yan S, Wang SS, Wang CY. The correctness on model results with uncertainty in the objective function and the constraint set. The 15th European Academic Research Conference on Global Business, Economics, Finance and Banking, Zurich, Switzerland 2019.
[35] Kline SJ, McClintock FA. Describing uncertainties in single-sample experiments. Mech Eng 1953;75:3−8.
[36] Garey MR, Johnson DS. Computers and intractability: a guide to the theory of NP-completeness. San Francisco, CA: W.H. Freeman & Company; 1979.